


JOINT INVENTORS

"EXPRESS MAIL" mailing label No.
ET924102454US.

Date of Deposit: January 8, 2002

I hereby certify that this paper (or fee) is
being deposited with the United States Postal
Service "EXPRESS MAIL POST OFFICE TO
ADDRESSEE" service under 37 CFR §1.10 on
the date indicated above and is addressed to:
Commissioner for Patents, Washington, D.C.
20231


Richard Zimmermann

APPLICATION FOR UNITED STATES LETTERS PATENT

SPECIFICATION

TO ALL WHOM IT MAY CONCERN:

Be it known that we, Lutz Gissmann, a citizen of Germany, residing
at Piroweeg 1, 69168 Wiesloch, in the Country of Germany, and Martin Müller,
a citizen of Germany, residing at 1351 N. Hoyne, Chicago 60622 and State of
Illinois have invented a new and useful PAPILLOMA VIRUS CAPSOMERE
VACCINE FORMULATIONS AND METHODS OF USE, of which the following is
a specification.

New Patent Application for:

Gissmann, *et al.*

For:

Papilloma Virus Capsomere Vaccine
Formulations and Methods of Use

Mailing Certification for:

Patent Application

Attorney Docket No:

27013/38150

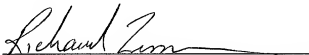
"EXPRESS MAIL" mailing label No.:

ET924102454US

Date of Deposit:

January 8, 2002

I hereby certify that this paper (or fee) is being deposited with the United States Postal Service "EXPRESS MAIL POST OFFICE TO ADDRESSEE" service under 37 CFR §1.10 on the date indicated above and is addressed to the Commissioner for Patents, Washington, D.C., 20231.


Richard Zimmermann

PAPILLOMA VIRUS CAPSOMERE VACCINE FORMULATIONS AND METHODS OF USE

FIELD OF THE INVENTION

The present invention relates to vaccine formulations
5 comprising papilloma virus proteins, either as fusion proteins, truncated
proteins, or truncated fusion proteins. The invention further embraces
methods for producing capsomeres of the formulations, as well as
prophylactic and therapeutic methods for their use.

BACKGROUND

10 Infections with certain high-risk strains of genital papilloma
viruses in humans (HPV) -- for example, HPV 16, 18, or 45 -- are
believed to be the main risk factor for the formation of malignant tumors of
the anogenital tract. Of the possible malignancies, cervical carcinoma is by
far the most frequent: according to an estimate by the World Health
15 Organization (WHO), almost 500,000 new cases of the disease occur
annually. Because of the frequency with which this pathology occurs, the
connection between HPV infection and cervical carcinoma has been
extensively examined, leading to numerous generalizations.

For example, precursor lesions of cervical intraepithelial
20 neoplasia (CIN) are known to be caused by papilloma virus infections
[Crum, *New Eng. J. Med.* 310:880-883 (1984)]. DNA from the genomes
of certain HPV types, including for example, strains 16, 18, 33, 35, and
45, have been detected in more than 95% of tumor biopsies from patients
with this disorder, as well as in primary cell lines cultured from the
25 tumors. Approximately 50 to 70% of the biopsied CIN tumor cells have
been found to include DNA derived only from HPV 16.

The protein products of the HPV 16 and HPV 18 early genes
E6 and E7 have been detected in cervical carcinoma cell lines as well as in

human keratinocytes transformed *in vitro* [Wettstein, *et al.*, in PAPILLOMA VIRUSES AND HUMAN CANCER, Pfister (Ed.), CRC Press: Boca Raton, FL 1990 pp 155-179] and a significant percentage of patients with cervical carcinoma have anti-E6 or anti-E7 antibodies. The E6 and E7 proteins have been shown to participate in induction of cellular DNA synthesis in human cells, transformation of human keratinocytes and other cell types, and tumor formation in transgenic mice [Arbelt, *et al.*, *J. Virol.*, 68:4358-4364 (1994); Auewarakul, *et al.*, *Mol. Cell. Biol.* 14:8250-8258 (1994); Barbosa, *et al.*, *J. Virol.* 65:292-298 (1991); Kaur, *et al.*, *J. Gen. Virol.* 70:1261-1266 (1989); Schlegel, *et al.*, *EMBO J.*, 7:3181-3187 (1988)]. The constitutive expression of the E6/E7 proteins appears to be necessary to maintain the transformed condition of HPV-positive tumors.

Despite the capacity of some HPV strains to induce neoplastic phenotypes *in vivo* and *in vitro*, still other HPV types cause benign genital warts such as condylomata acuminata and are only rarely associated with malignant tumors [Ikenberg. In Gross, *et al.*, (eds.) GENITAL PAPILLOMAVIRUS INFECTIONS. Springer Verlag: Berlin, pp., 87-112]. Low risk strains of this type include, for example, HPV 6 and 11.

Most often, genital papilloma viruses are transmitted between humans during intercourse which in many instances leads to persistent infection in the anogenital mucous membrane. While this observation suggests that either the primary infection induces an inadequate immune response or that the virus has developed the ability to avoid immune surveillance, other observations suggest that the immune system is active during primary manifestation as well as during malignant progression of papilloma virus infections [Altmann *et al.* in VIRUSES AND CANCER, Minson *et al.*, (eds.) Cambridge University Press, (1994) pp. 71-80].

For example, the clinical manifestation of primary infection by rabbit and bovine papilloma virus can be prevented by vaccination with wart extracts or viral structural proteins [Altmann, *et al.*, *supra*; Campo,

Curr. Top. In Microbiol and Immunol. 186:255-266 (1994); Yindle and Frazer, *Curr. Top. In Microbiol. and Immunol.* 186:217-253 (1994)]. Rodents previously vaccinated with vaccinia recombinants encoding HPV 16 early proteins E6 or E7, or with synthetic E6 or E7 peptides, are similarly protected from tumor formation after inoculation of HPV 16 transformed autologous cells [Altman, *et al.*, *supra*; Campo, *et al.*, *supra*; Yindle and Frazer, *et al. supra*]. Regression of warts can be induced by the transfer of lymphocytes from regressor animals following infection by animal papilloma viruses. Finally, in immunosuppressed patients, such as, for example, recipients of organ transplants or individuals infected with HIV, the incidence of genital warts, CIN, and anogenital cancer is elevated.

To date, no HPV vaccinations have been described which comprise human papilloma virus late L1 protein in the form of capsomeres which are suitable both for prophylactic and therapeutic purposes. Since the L1 protein is not present in malignant genital lesions, vaccination with L1 protein does not have any therapeutic potential for these patients. Construction of chimeric proteins, comprising amino acid residues from L1 protein and, for example E6 or E7 protein, which give rise to chimeric capsomeres, combines prophylactic and therapeutic functions of a vaccine. A method for high level production of chimeric capsomeres would therefore be particularly desirable, in view of the possible advantages offered by such a vaccine for prophylactic and therapeutic intervention.

Thus there exists a need in the art to provide vaccine formulations which can prevent or treat HPV infection. Methods to produce vaccine formulations which overcome problems known in the art to be associated with recombinant HPV protein expression and purification would manifestly be useful to treat the population of individuals already infected with HPV as well as useful to immunize the population of individuals susceptible to HPV infection.

SUMMARY OF THE INVENTION

The present invention provides therapeutic and prophylactic vaccine formulations comprising chimeric human papilloma capsomeres.

The invention also provides therapeutic methods for treating patients

- 5 infected with an HPV as well as prophylactic methods for preventing HPV infection in a susceptible individual. Methods for production and purification of capsomeres and proteins of the invention are also contemplated.

- 10 In one aspect of the invention, prophylactic vaccinations for prevention of HPV infection are considered which incorporate the structural proteins L1 and L2 of the papilloma virus. Development of a vaccine of this type faces significant obstacles because papilloma viruses cannot be propagated to adequate titers in cell cultures or other experimental systems to provide the viral proteins in sufficient quantity for economical vaccine
- 15 production. Moreover, recombinant methodologies to express the proteins are not always straightforward and often results in low protein yield. Recently, virus-like particles (VLPs), similar in make up to viral capsid structures, have been described which are formed in Sf-9 insect cells upon expression of the viral proteins L1 and L2 (or L1 on its own) using
- 20 recombinant vaccinia or baculovirus. Purification of the VLPs can be achieved very simply by means of centrifugation in CsCl or sucrose gradients [Kimbauer, *et al.*, *Proc. Natl. Acad. Sci. (USA)*, 99:12180-12814 (1992); Kimbauer, *et al.*, *J. Virol.* 67:6929-6936 (1994); Proso, *et al.*, *J. Virol.* 67:14:1936-1944 (1992); Sasagawa, *et al.*, *Virology* 2016:126-195 (1995); Volpers, *et al.*, *J. Virol.* 69:3258-3264 (1995); Zhou, *et al.*, *J. Gen. Virol.* 74:762-769 (1993); Zhou, *et al.*, *Virology* 185:251-257 (1991)]. WO 93/02184 describes a method in which papilloma virus-like particles (VLPs) are used for diagnostic applications or as a vaccine against infections caused by the papilloma virus. WO 94/00152 describes

recombinant production of L1 protein which mimics the conformational neutralizing epitope on human and animal papilloma virions.

In another aspect of the invention, therapeutic vaccinations are provided to relieve complications of, for example, cervical carcinoma or precursor lesions resulting from papilloma virus infection, and thus represent an alternative to prophylactic intervention. Vaccinations of this type may comprise early papilloma virus proteins, principally E6 or E7, which are expressed in the persistently infected cells. It is assumed that, following administration of a vaccination of this type, cytotoxic T-cells might be activated against persistently infected cells in genital lesions. The target population for therapeutic intervention is patients with HPV-associated pre-malignant or malignant genital lesions. PCT patent application WO 93/20844 discloses that the early protein E7 and antigenic fragments thereof of the papilloma virus from HPV or BPV is therapeutically effective in the regression, but not in the prevention, of papilloma virus tumors in mammals. While early HPV proteins have been produced by recombinant expression in *E. coli* or suitable eukaryotic cell types, purification of the recombinant proteins has proven difficult due to inherent low solubility and complex purification procedures which generally require a combination of steps, including ion exchange chromatography, gel filtration and affinity chromatography.

According to the present invention, vaccine formulations comprising papilloma virus capsomeres are provided which comprise either: (i) a first protein that is an intact viral protein expressed as a fusion protein comprised in part of amino acid residues from a second protein; (ii) a truncated viral protein; (iii) a truncated viral protein expressed as a fusion protein comprised in part of amino acid residues from a second protein, or (iv) some combination of the three types of proteins. According to the invention, vaccine formulations are provided comprising capsomeres of bovine papilloma virus (BPV) and human papilloma virus. Preferred

bovine virus capsomeres comprise protein from bovine papilloma virus type

I. Preferred human virus capsomeres comprise proteins from any one of human papilloma virus strains HPV6, HPV11, HPV16, HPV18, HPV33, HPV35, and HPV45. The most preferred vaccine formulations comprise capsomeres comprising proteins from HPV16.

In one aspect, capsomere vaccine formulations of the invention comprise a first intact viral protein expressed as a fusion protein with additional amino acid residues from a second protein. Preferred intact viral proteins are the structural papilloma viral proteins L1 and L2.

Capsomeres comprised of intact viral protein fusions may be produced using the L1 and L2 proteins together or the L1 protein alone. Preferred capsomeres are made up entirely of L1 fusion proteins, the amino acid sequence of which is set out in SEQ ID NO: 2 and encoded by the polynucleotide sequence of SEQ ID NO: 1. Amino acids of the second protein can be derived from numerous sources (including amino acid residues from the first protein) as long as the addition of the second protein amino acid residues to the first protein permits formation of capsomeres.

Preferably, addition of the second protein amino acid residues inhibits the ability of the intact viral protein to form virus-like particle structures; most preferably, the second protein amino acid residues promote capsomere formation. In one embodiment of the invention, the second protein may be any human tumor antigen, viral antigen, or bacterial antigen which is important in stimulating an immune response in neoplastic or infectious disease states. In a preferred embodiment, the second protein is also a papilloma virus protein. It also preferred that the second protein be the expression product of papilloma virus early gene. It is also preferred, however, that the second protein be selected from group of E1, E2, E3, E4, E5, E6, and E7 -- early gene products encoded in the genome of papilloma virus strains HPV6, HPV11, HPV18, HPV33, HPV35, or HPV45. It is most preferred that the second protein be encoded by the HPV16

E7 gene, the open reading frame of which is set out in SEQ ID NO: 3. Capsomeres assembled from fusion protein subunits are referred to herein as chimeric capsomeres. In one embodiment, the vaccine formulation of the invention is comprised of chimeric capsomeres wherein L1 protein amino acid residues make up approximately 50 to 99% of the total fusion protein amino acid residues. In another embodiment, L1 amino acid residues make up approximately 60 to 90% of the total fusion protein amino acid residues; in a particularly preferred embodiment, L1 amino acids comprise approximately 80% of the fusion protein amino acid residues.

In another aspect of the invention, capsomere vaccine formulations are provided that are comprised of truncated viral proteins having a deletion of one or more amino acid residues necessary for formation of a virus-like particle. It is preferred that the amino acid deletion not inhibit formation of capsomeres by the truncated protein, and it is most preferred that the deletion favor capsomere formation. Preferred vaccine formulations of this type include capsomeres comprised of truncated L1 with or without L2 viral proteins. Particularly preferred capsomeres are comprised of truncated L1 proteins. Truncated proteins contemplated by the invention include those having one or more amino acid residues deleted from the carboxy terminus of the protein, or one or more amino acid residues deleted from the amino terminus of the protein, or one or more amino acid residues deleted from an internal region (*i.e.*, not from either terminus) of the protein. Preferred capsomere vaccine formulations are comprised of proteins truncated at the carboxy terminus. In formulations including L1 protein derived from HPV16, it is preferred that from 1 to 34 carboxy terminal amino acid residues are deleted. Relatively shorter deletions are also contemplated which offer the advantage of minor modification of the antigenic properties of the L1 proteins and the capsomeres formed thereof. It is most preferred, however, that 34 amino

acid residues be deleted from the L1 sequence, corresponding to amino acids 472 to 505 in HPV16 set out in SEQ ID NO: 2, and encoded by the polynucleotide sequence corresponding to nucleotides 1414 to 1516 in the human HPV16 L1 coding sequence set out in SEQ ID NO: 1.

5 When a capsomere vaccine formulation is made up of proteins bearing an internal deletion, it is preferred that the deleted amino acid sequence comprise the nuclear localization region of the protein. In the L1 protein of HPV 16, the nuclear localization signal is found from about amino acid residue 499 to about amino acid residue 505. Following
10 expression of L1 proteins wherein the NLS has been deleted, assembly of capsomere structures occurs in the cytoplasm of the host cell. Consequently, purification of the capsomeres is possible from the cytoplasm instead of from the nucleus where intact L1 proteins assemble into capsomeres. Capsomeres which result from assembly of truncated
15 proteins wherein additional amino acid sequences do not replace the deleted protein sequences are necessarily not chimeric in nature.

 In still another aspect of the invention, capsomere vaccine formulations are provided comprising truncated viral protein expressed as a fusion protein adjacent amino acid residues from a second protein.
20 Preferred truncated viral proteins of the invention are the structural papilloma viral proteins L1 and L2. Capsomeres comprised of truncated viral protein fusions may be produced using L1 and L2 protein components together or L1 protein alone. Preferred capsomeres are those comprised of L1 protein amino acid residues. Truncated viral protein components of the
25 fusion proteins include those having one or more amino acid residues deleted from the carboxy terminus of the protein, or one or more amino acid residues deleted from the amino terminus of the protein, or one or more amino acid residues deleted from an internal region (*i.e.*, not from either terminus) of the protein. Preferred capsomere vaccine formulations
30 are comprised of proteins truncated at the carboxy terminus. In those

formulations including L1 protein derived from HPV16, it is preferred that from 1 to 34 carboxy terminal amino acid residues are deleted. Relatively shorter deletions are also contemplated that offer the advantage of minor modification of the antigenic properties of the L1 protein component of the fusion protein and the capsomeres formed thereof. It is most preferred, however, that 34 amino acid residues be deleted from the L1 sequence, corresponding to amino acids 472 to 505 in HPV16 set out in SEQ ID NO: 2, and encoded by the polynucleotide sequence corresponding to nucleotides 1414 to 1516 in the human HPV16 L1 coding sequence set out in SEQ ID NO: 1. When the vaccine formulation is comprised of capsomeres made up of proteins bearing an internal deletion, it is preferred that the deleted amino acid sequence comprise the nuclear localization region, or sequence, of the protein.

Amino acids of the second protein can be derived from numerous sources as long as the addition of the second protein amino acid residues to the first protein permits formation of capsomeres. Preferably, addition of the second protein amino acid residues promotes or favors capsomere formation. Amino acid residues of the second protein can be derived from numerous sources, including amino acid residues from the first protein. In a preferred embodiment, the second protein is also a papilloma virus protein. It also preferred that the second protein be the expression product of papilloma virus early gene. It is most preferred, however, that the second protein be selected from group of early gene products encoding by papilloma virus E1, E2, E3, E4, E5, E6, and E7 genes. In one embodiment, the vaccine formulation of the invention is comprised of chimeric capsomeres wherein L1 protein amino acid residues make up approximately 50 to 99% of the total fusion protein amino acid residues. In another embodiment, L1 amino acid residues make up approximately 60 to 90% of the total fusion protein amino acid residues; in

a particularly preferred embodiment, L1 amino acids comprise approximately 80% of the fusion protein amino acid residues.

In a preferred embodiment of the invention, proteins of the vaccine formulations are produced by recombinant methodologies, but in formulations comprising intact viral protein, the proteins may be isolated from natural sources. Intact proteins isolated from natural sources may be modified *in vitro* to include additional amino acid residues to provide a fusion protein of the invention using covalent modification techniques well known and routinely practiced in the art. Similarly, in formulations comprising truncated viral proteins, the proteins may be isolated from natural sources as intact proteins and hydrolyzed *in vitro* using chemical hydrolysis or enzymatic digestion with any of a number of site-specific or general proteases, the truncated protein subsequently modified to include additional amino acid residues as described above to provide a truncated fusion protein of the invention.

In producing capsomeres, recombinant molecular biology techniques can be utilized to produce DNA encoding either the desired intact protein, the truncated protein, or the truncated fusion protein. Recombinant methodologies required to produce a DNA encoding a desired protein are well known and routinely practiced in the art. Laboratory manuals, for example Sambrook, *et al.*, (eds.), MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Harbor Press: Cold Spring Harbor, NY (1989) and Ausubel *et al.*, (eds.), PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. (1994-1997), describe in detail techniques necessary to carry out the required DNA manipulations. For large-scale production of chimeric capsomeres, protein expression can be carried out using either viral or eukaryotic vectors. Preferable vectors include any of the well known prokaryotic expression vectors, recombinant baculoviruses, COS cell specific vectors, vaccinia recombinants, or yeast-specific expression constructs. When recombinant proteins are used to

provide capsomeres of the invention, the proteins may first be isolated from the host cell of its expression and thereafter incubated under conditions which permit self-assembly to provide capsomeres. Alternatively, the proteins may be expressed under conditions wherein capsomeres are formed in the host cell.

The invention also contemplates processes for producing capsomeres of the vaccine formulations. In one method, L1 proteins are expressed from DNA encoding six additional histidines at the carboxy terminus of the L1 protein coding sequence. L1 proteins expressed with additional histidines (His L1 proteins) are most preferably expressed in *E. coli* and the His L1 proteins can be purified using nickel affinity chromatography. His L1 proteins in cell lysate are suspended in a denaturation buffer, for example, 6 M guanidine hydrochloride or a buffer of equivalent denaturing capacity, and then subjected to nickel chromatography. Protein eluted from the nickel chromatography step is renatured, for example in 150 mM NaCl, 1 mM CaCl_2 , 0.01 % Triton-X 100, 10 mM HEPES (N-2-hydroxyethyl piperazine-N'-2 ethane sulfonic acid), pH 7.4. According to a preferred method of the invention, assembly of capsomeres takes place after dialysis of the purified proteins, preferably after dialysis against 150 mM NaCl, 25 mM Ca^{2+} , 10 % DMSO (dimethyl sulfoxide), 0.1 % Triton-X 100, 10 mM Tris [tris-(hydroxymethyl) amino-methane] acetic acid with a pH value of 5.0.

Formation of capsomeres can be monitored by electron microscopy, and, in instances wherein capsomeres are comprised of fusion proteins, the presence of various protein components in the assembled capsomere can be confirmed by Western blot analysis using specific antisera.

According to the present invention, methods are provided for therapeutic treatment of individuals infected with HPV comprising the step of administering to a patient in need thereof an amount of a vaccine

formulation of the invention effective to reduce the level of HPV infection. The invention also provide methods for prophylactic treatment of individuals susceptible to HPV infection comprising the step of administering to an individual susceptible to HPV infection an amount of a vaccine formulation of the invention effective to prevent HPV infection.

While infected individuals can be easily identified using standard diagnostic techniques, susceptible individuals may be identified, for example, as those engaged in sexual relations with an infected individual. However, due to the high frequency of HPV infection, all sexually active persons are susceptible to papilloma virus infection.

Administration of a vaccine formulation can include one or more additional components such as pharmaceutically acceptable carriers, diluents, adjuvants, and/or buffers. Vaccines may be administered at a single time or at multiple times. Vaccine formulation of the invention may be delivered by various routes including, for example, oral, intravenous, intramuscular, nasal, rectal, transdermal, vaginal, subcutaneous, and intraperitoneal administration.

Vaccine formulations of the invention offer numerous advantages when compared to conventional vaccine preparations. As part of a therapeutic vaccination, capsomeres can promote elimination of persistently infected cells in, for example, patients with CIN or cervical carcinoma. Additionally, therapeutic vaccinations of this type can also serve a prophylactic purpose in protecting patients with CIN lesions from re-infection. As an additional advantage, capsomeres can escape neutralization by pre-existing anticapsid antibodies and thereby possess longer circulating half-life as compared to chimeric virus-like particles.

Vaccine formulations comprising chimeric capsomeres can provide the additional advantage of increased antigenicity of both protein components of the fusion protein from which the capsomere is formed.

For example, in a VLP, protein components of the underlying capsomere

may be buried in the overall structure as a result of internalized positioning within the VLP itself. Similarly, epitopes of the protein components may be sterically obstructed as a result of capsomere-to-capsomere contact, and therefore inaccessible for eliciting an immune response. Preliminary results using L1/E7 fusion proteins to produce VLPs support this position in that no antibody response was detected against the E7 component. This observation is consistent with previous results which indicate that the carboxy terminal region of L1 forms inter-pentameric arm structures that allow assembly of capsomeres into capsids [Garcia, *et al.*, *J. Virol.* 71: 2988-2995 (1997)]. Presumably in a chimeric capsomere structure, both protein components of the fusion protein substructure are accessible to evoke an immune response. Capsomere vaccines would therefore offer the additional advantage of increased antigenicity against any protein component, including, for example, neutralizing epitopes from other virus proteins, expressed as a fusion with L1 amino acid sequences.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is illustrated by the following examples. Example 1 describes construction of expression vectors to produce fusion, or chimeric, viral proteins. Example 2 relates to generation of recombinant baculoviruses for expression of viral proteins. Example 3 addresses purification of capsomeres. Example 4 describes an immunization protocol for production of antisera and monoclonal antibodies. Example 5 provides a peptide ELISA to quantitate capsomere formation. Example 6 describes an antigen capture ELISA to quantitate capsomere formation. Example 7 provides a hemagglutinin assay to assay for the induction of neutralizing antibodies.

Example 1 Construction of Chimeric L1 Genes

DNA encoding the HPV 16 L1 open reading frame was excised from plasmid 16-114/k-L1/L2-pSynxtVT [Kirnbauer *et al.*, *J. Virol.*

5 67:6929-6936 (1994)] using *Bgl*III and the resulting fragment subcloned into pUC19 (New England Biolabs, Beverly, MA) previously linearized at the unique *Bam*HI restriction site. Two basic expression constructs were first generated to permit subsequent insertion of DNA to allow fusion protein expression. One construct encoded HPV 16 L1 Δ 310 having a nine amino
10 acid deletion; the deleted region was known to show low level homology with all other papilloma virus L1 proteins. The second construct, HPV 16 L1 Δ C, encoded a protein having a 34 amino acid deletion of the carboxy terminal L1 residues. Other constructs include an *Eco*RV restriction site at the position of the deletion for facilitated insertion of DNA encoding other
15 protein sequences. Addition of the *Eco*RV site encodes two non-L1 protein amino acids, aspartate and isoleucine.

A. Generation of an HPV 16 L1 Δ 310 expression construct

Two primers (SEQ ID NOs: 5 and 6) were designed to amplify the pUC19 vector and the complete HPV 16 L1 coding sequence,
20 except nucleotides 916 through 942 in SEQ ID NO: 1. Primers were synthesized to also introduce a unique *Eco*RV restriction site (underlined in SEQ ID NOs: 5 and 6) at the termini of the amplification product.

CCCCGATATCGCCTTTAATGTATAAATCGTCTGG
SEQ ID NO: 5

25 CCCCGATATCTCAAATTATTTTCTACACCTAGTG
SEQ ID NO: 6

The resulting PCR product was digested with *Eco*RV to provide complementary ends and the digestion product circularized by ligation.

Ligated DNA was transformed into *E. coli* using standard techniques and plasmids from resulting colonies were screened for the presence of an *EcoRV* restriction site. One clone designated HPV 16 L1 Δ 310 was identified as having the appropriate twenty-seven nucleotide deletion and this construct was used to insert DNA fragments encoding other HPV 16 proteins at the *EcoRV* site as discussed below.

B. Generation of an HPV 16 L1 Δ C expression constructs

Two primers (SEQ ID NOs: 7 and 8) were designed complementary to the HPV 16 L1 open reading frame such that the primers abutted each other to permit amplification in reverse directions on the template DNA comprising HPV 16 L1-encoding sequences in pUC19 described above.

AAAGATATCTTGTAGTAAAAATTGCGTCCTAAAGGAAAC
SEQ ID NO: 7

AAAGATATCTAATCTACCTCTACAACGCTAAACGCAAAAAACG
SEQ ID NO: 8

Each primer introduced an *EcoRV* restriction site at the terminus of the amplification product. In the downstream primer (SEQ ID NO: 8), the *EcoRV* site was followed by a TAA translational stop codon positioned such that the amplification product, upon ligation of the *EcoRV* ends to circularize, would include deletion of the 34 carboxy terminal L1 amino acids. PCR was performed to amplify the partial L1 open reading frame and the complete vector. The amplification product was cleaved with *EcoRV*, circularized with T4 DNA ligase, and transformed into *E. coli* DH5 α cells. Plasmids from viable clones were analyzed for the presence of an *EcoRV* site which would linearize the plasmid. One positive

construct designated pUCHPV16L1ΔC was identified and used to insert DNA from other HPV 16 proteins utilizing the *EcoRV* site.

C. Insertion of DNA fragments into HPV 16 L1 Δ310 and HPV16L1ΔC

DNA fragments of HPV 16 E7 encoding amino acids 1-50, 1-60, 1-98, 25-75, 40-98, 50-98 in SEQ ID NO: 4 were amplified using primers that introduced terminal 5' *EcoRV* restriction sites in order to facilitate insertion of the fragment into either HPV 16 L1 Δ310 and HPV16L1ΔC modified sequence. In the various amplification reactions, primer E7.1 (SEQ ID NO: 9) was used in combination with primer E7.2 (SEQ ID NO: 10) to generate a DNA fragment encoding E7 amino acids 1-50; with primer E7.3 (SEQ ID NO: 11) generate a DNA fragment encoding E7 amino acids 1-60; or with primer E7.4 (SEQ ID NO: 12) generate a DNA fragment encoding E7 amino acids 1-98. In other amplification reactions, primer pairs E7.5 (SEQ ID NO: 13) and E7.6 (SEQ ID NO: 14) were used to amplify a DNA fragment encoding E7 amino acids 25-75; E7.7 (SEQ ID NO: 15) and E7.4 (SEQ ID NO: 12) were used to amplify a DNA fragment encoding E7 amino acids 40-98; and E7.8 (SEQ ID NO: 16) and E7.4 (SEQ ID NO: 12) were used to amplify a DNA fragment encoding E7 amino acids 50-98.

20 Primer E7.1 SEQ ID NO: 9
AAAAGATATCATGCATGGAGATACACCTACATTGC

Primer E7.2 SEQ ID NO: 10
TTTTGATATCGGCTCTGTCCGGTTCTGCTTGTC

25 Primer E7.3 SEQ ID NO: 11
TTTTGATATCCTTGCAACAAAAGGTTACAATATTGTAATGGGCC

Primer E7.4 SEQ ID NO: 12
AAAAGATATCTGGTTTCTGAGAACAGATGGGGCAC

Primer E7.5 SEQ ID NO: 13
TTTTGATATCGATTATGAGCAATTAAATGACAGCTCAC

5 Primer E7.6 SEQ ID NO: 14
TTTTGATATCGTCTACGTGTGTGCTTTGTACGCAC

Primer E7.7 SEQ ID NO: 15
TTTATCGATATCGGTCCAGCTGGACAAGCAGAACCGGAC

10 Primer E7.8 SEQ ID NO: 16
TTTTGATATCGATGCCCCATTACAATATTGTAACCTTTTG

Similarly, nucleotides from DNA encoding the influenza matrix protein (SEQ ID NO: 17) was amplified using the primer pair set out in SEQ ID NOs: 19 and 20. Both primers introduced an *EcoRV* restriction site in the amplification product.

15 TTTTGATATCGATATGGAATGGCTAAAGACAAGACCAATC
SEQ ID NO: 19

TTTTGATATCGTTGTTTGGATCCCCATTCCCATTG
SEQ ID NO: 20

20 PCR products from each amplification reaction were cleaved with *EcoRV* and inserted into the *EcoRV* site of either the HPV 16 L1 Δ 310 and HPV16L1 Δ C sequences previously linearized with the same enzyme. In order to determine the orientation of inserts in plasmids encoding E7 amino acids 25-75 and 50-98 and plasmid including influenza matrix protein, *ClaI* digestion was employed, taking advantage of a



10

15

20

Primer P2 SEQ ID NO: 22

5

Following the first two amplification reactions, the two purified products were used as templates in another PCR reaction using primers P1 and P4 only. The resulting amplification product was digested with enzymes *Eco*NI and *Hind*III inserted into the HPV 16 L1ΔC expression construct described above following digestion with the same enzymes. The resulting expression construct differed from the original HPV16L1ΔC construct with DNA encoding L1 and E7 amino acids 1-50 by loss of the two internal *Eco*RV restriction sites. The first *Eco*RV site was replaced by DNA encoding native L1 alanine and glycine amino acids in this position and the second was replaced by a translational stop signal. In addition, the expression construct, designated HPV 16 L1ΔC* E7 1-52, contained the first 52 amino acids of HPV 16 E7 as a result of using primer P4 which also encodes E7 amino acids residues histidine at position 51 and tyrosine at position 52. HPV 16 L1ΔC* E7 1-52 was then used to generate additional HPV 16 L1ΔC expression constructs further including DNA encoding E7 amino acids 1-55 using primer P1 (SEQ ID NO: 21) in combination with primer P5 (SEQ ID NO: 25), E7 amino acids 1-60 with primer pair P1 and P6 (SEQ ID NO: 26), and E7 amino acids 1-65 with primer pair P1 and P7 (SEQ ID NO: 27). The additional amino acid-encoding DNA sequences in the amplification products arose from design of the primers to include additional nucleotides for the desired amino acids.

Primer P5 SEQ ID NO: 25
CATCTGAAGCTTAACAATATTGTAATGGGCTCTGTCCG

Primer P6 SEQ ID NO: 26
CATCTGAAGCTTACTTGCAACAAAAGGTTA-
CAATATTGTAATGGGCTCTGTCCG

Primer P7 SEQ ID NO: 27
CATCTGAAGCTTAAAGCGTAGAGTCACACTTGCAAC-
AAAAGGTTACAATATTGTAATGGGCTCTGTCCG

Similarly, HPV 16 L1ΔC* E7 1-70 was generated using template DNA
encoding HPV 16 L1ΔC* E7 1-66 and the primer pair P1 and P8 (SEQ ID
NO: 28).

Primer P8 SEQ ID NO: 28
CATCTGAAGCTTATTGTACGCACAAC-
CGAAGCGTAGAGTCACACTTG

Following each PCR reaction, the amplification products were digested
with *Eco*NI and *Hind*III and inserted into HPV16L1ΔC previously digested
with the same enzymes. Sequences of each constructs were determined
using an Applied Biosystems Prism 377 sequencing instrument with
fluorescent chain terminating dideoxynucleotides [Prober *et al.*, *Science*
238:336-341 (1987)].

Example 2 Generation of Recombinant Baculoviruses

Spodoptera frugiperda (Sf9) cells were grown in suspension or
monolayer cultures at 27° in TNMFH medium (Sigma) supplemented with
10% fetal calf serum and 2 mM glutamine. For HPV 16 L1-based
recombinant baculovirus construction, Sf9 cells were transfected with 10 μg
of transfer plasmid together with 2 μg of linearized Baculo-Gold DNA

(PharMingen, San Diego, CA). Recombinant viruses were purified by according to manufacturer's suggested protocol.

To test for expression of HPV 16 L1 protein, 10^5 Sf9 cells were infected with baculovirus recombinant at a multiplicity of infection (m.o.i) of 5 to 10. After incubation for three to four days at 28°C, media was removed and cells were washed with PBS. The cells were lysed in SDS sample buffer and analyzed by SDS-PAGE and Western blotting using anti-HPV16 L1 and anti-HPV16 E7 antibodies.

In order to determine which of the chimeric L1 protein expression constructs would preferentially produce capsomeres, extracts from transfected cells were subjected to gradient centrifugation. Fractions obtained from the gradient were analyzed for L1 protein content by Western blotting and for VLP formation by electron microscopy. The results are shown in Table 1.

The intact HPV L1 protein, as well as the expression products HPV 16 L1Δ310 and HPV 16 L1ΔC, each were shown to produce capsomeres and virus-like particles in equal proportions. When E7 coding sequences were inserted into the HPV 16 L1Δ310 vector, only fusion proteins including E7 amino acids 1 to 50 produced gave rise to detectable capsomere formation.

When E7 encoding DNA was inserted into the HPV 16 L1ΔC vector, all fusion proteins were found to produce capsomeres; chimeric proteins including E7 amino acid residues 40-98 produced the highest level of exclusively capsomere structures. Chimeric proteins including E7 amino acids 1-98 and 25-75 both produced predominantly capsomeres, even thorough virus-like particle formation was also observed. The chimeric protein including E7 amino acids 1-60 resulted in nearly equal levels of capsomere and virus-like particle production.

When E7 sequences were inserted into the HPV 16 L1ΔC vector, all fusion proteins were shown to produce capsomeres. Insertion of

DNA encoding E7 residues 1-52, 1-55, and 1-60 produced the highest level of capsomeres, but equal levels of virus-like particle production were observed. While insertion of DNA encoding E7 DNA for residues 1-65, 1-70, 25-75, 40-98, and 1-98 resulted in comparatively lower levels or undetectable levels of capsid, capsomeres were produced in high quantities.

TABLE 1
Capsomere and Capsid Forming Capacity of
Chimeric HPV L1 Proteins

	<u>L1 Expression Construct</u>	<u>Insert</u>	<u>Capsomere Yield</u>	<u>Capsid Yield</u>
10	HVP 16 L1	None	+++++	+++++
	HPV 16 L1Δ310	None	+++	++
	HPV 16 L1ΔC	None	++++	++++
	HPV 16 L1Δ310	E7 1-98	-	-
15	HPV 16 L1Δ310	E7 1-50	++	-
	HPV 16 L1Δ310	E7 25-75	-	-
	HPV 16 L1Δ310	E7 50-98	-	-
	HPV 16 L1ΔC	E7 1-98	+++	+
	HPV 16 L1ΔC	E7 25-75	+++	+
20	HPV 16 L1ΔC	E7 50-98	+	+
	HPV 16 L1ΔC	E7 1-60	+++++	+++++
	HPV 16 L1ΔC	E7 40-98	++++	-
	HPV 16 L1ΔC	Influenza	+++	+
	HPV 16 L1Δ*C	E7 1-52	+++++	+++++
25	HPV 16 L1Δ*C	E7 1-55	+++++	+++++
	HPV 16 L1Δ*C	E7 1-60	+++	++++
	HPV 16 L1Δ*C	E7 1-65	++	-
	HPV 16L1Δ*C	E7 1-70	++	-

Example 3 Purification of Capsomeres

Trichopulsia ni (TN) High Five cells were grown to a density of approximately 2×10^6 cells/ml in Ex-Cell 405 serum-free medium (JRH Biosciences). Approximately 2×10^8 cells were pelleted by centrifugation at 1000 x g for 15 minutes, resuspended in 20 ml of medium, and infected with recombinant baculoviruses at m.o.i of 2 to 5 for 1 hour at room temperature. After addition of 200 ml medium, cells were plated and incubated for 3 to 4 days at 27°C. Following incubation, cells were harvested, pelleted, and resuspended in 10 ml of extraction buffer.

The following steps were performed at 4°C. Cells were sonicated for 45 seconds at 60 watts and the resulting cell lysate was centrifuged at 10,000 rpm in a Sorval SS34 rotor. The supernatant was removed and retained while the resulting pellet was resuspended in 6 ml of extraction buffer, sonicated for an additional 3 seconds at 60 watts, and centrifuged again. The two supernatants were combined, layered onto a two-step gradient containing 14 ml of 40% sucrose on top of 8 ml of CsCl solution (4.6 g CsCl per 8 ml in extraction buffer), and centrifuged in a Sorval AH629 swinging bucket rotor for 2 hours at 27,000 rpm at 10°C. The interface region between the CsCl and the sucrose along with the CsCl complete layer were collected into 13.4 ml Quickseal tubes (Beckman) and extraction buffer added to adjust the volume 13.4 ml. Samples were centrifuged overnight at 50,000 rpm at 20°C in a Beckman 70 TI rotor. Gradients were fractionated (1 ml per fraction) by puncturing tubes on top and bottom with a 21-gauge needle. Fractions were collected from each tube and 2.5 µl of each fraction were analyzed by a 10% SDS-polyacrylamide gel and Western blotting using an anti-HPV16 L1 antibody.

Virus-like particles and capsomeres were separated from the fractions identified above by sedimentation on 10 to 50% sucrose gradients. Peak fractions from CsCl gradients were pooled and dialyzed for 2 hours against 5 mM HEPES (pH 7.5). Half of the dialysate was used to produce capsomeres by disassembly of intact VLPs overnight by adding EDTA (final concentration 50 mM), EGTA (50

mM), DTT (30 mM), NaCl (100 mM), and Tris/HCl, pH 8.0, (10 mM). As control, NaCl and Tris/HCl only were added to the other half.

For analysis of capsomeres produced from disassembled VLPs, EDTA, EGTA, and DTT (final concentration 5 mM each) were added to the sucrose cushions which were centrifuged at 250,000 x g for 2 to 4 hours at 4°C. Fractions were collected by puncturing tubes from the bottom. A 1:10 dilution of each fraction was then analyzed by antigen capture ELISA.

Example 4 Immunization Protocol for Production of Polyclonal Antisera and Monoclonal Antibodies

Balb/c mice are immunized subcutaneously three times, every four weeks with approximately 60 µg of HPV chimeric capsomeres mixed 1:1 with complete or incomplete Freund's Adjuvants in a total volume of 100 µl. Six weeks after the third immunization, mice are sacrificed and blood is collected by cardiac puncture.

Example 5 Peptide ELISA to Quantitate Capsomere Formation

Microtiter plates (Dynatech) are coated overnight with 50 µl of peptide E701 [Muller *et al.*, 1982] at a concentration of 10 µg/ml in PBS. Wells are blocked for 2 hour at 37°C with 100 µl of buffer containing 5% BSA and 0.05% Tween 20 in PBS and washed three times with PBS containing 0.05% Tween 20. After the third wash, 50 µl of sera diluted 1:5000 in BSA/Tween 20/PBS is added to each well and incubation carried out for 1 hour. Plates are washed again as before and 50 µl of goat-anti-mouse peroxidase conjugate is added at a 1:5000 dilution. After 1 hour, plates are washed and stained using ABTS substrate (0.2 mg/ml, 2,2'-Azino-bis(3-ethylbenziazoline-β-sulfonic acid in 0.1 M Na-Acetate-Phosphate buffer (pH 4.2) with 4 µl 30% H₂O₂ per 10 ml). Extinction is measured after 1 hour at 490 nm in a Dynatech automated plate reader.

Example 6 Antigen Capture ELISA to Quantitate Capsomere Formation

To allow relative quantification of virus-like particles and capsomeres in fractions of CsCl gradients, an antigen capture ELISA was utilized.

5 Microtiter plates were coated overnight with 50 μ l/well of a 1:500 dilution (final concentration of 2 μ g per ml, in PBS) with a protein A purified mouse monoclonal antibody immunospecific for HPV 16 L1 (antibodies 25/C, MM07 and Ritti 1 were obtained from mice immunized with HPV 16 VLPs). Plates were blocked with 5% milk/PBS for 1 hour and 50 μ l of fractions of CsCl gradients were added.

10 for 1 hour at 37°C using a 1:300 dilution (in 5% milk/PBS). After three washings with PBS/0.05% Tween 20, 50 μ l of a polyclonal rabbit antiserum (1:3000 dilution in milk/PBS), raised against HPV 16 VLPs was added and plates were incubated at 37° for 1 hour. Plates were washed again and further incubated with 50 μ l of a goat-anti-rabbit peroxidase conjugate (Sigma) diluted 1:5000 in PBS

15 containing 5% milk for 1 hour. After final washing, plates were stained with ABTS substrate for 30 minutes and extinction measured at 490 nm in a Dynatech automated plate reader. As a negative control, the assay also included wells coated only with PBS.

To test monoclonal antibodies for capsomere specificity, VLPs with

20 EDTA/DTT to disassemble particles. Treated particle preparations were assayed in the antigen-capture ELISA and readings compared to untreated controls. For disassembly, 40 μ l of VLPs was incubated overnight at 4°C in 500 μ l of disruption buffer containing 30 mM DTT, 50 mM EGTA, 60 mM EDTA, 100 mM NaCl, and 100 mM Tris/HCl, pH 8.0. Aliquots of treated and untreated

25 particles were used in the above capture ELISA in a 1:20-1:40 dilution.

Example 7 Hemagglutinin Inhibition Assay

In order to determine the extent to which chimeric capsomere vaccines evoke production of neutralizing antibodies, a hemagglutination inhibition

30 assay is carried out as briefly described below. This assay is based on previous

observations that virus-like particles are capable of hemagglutinating red blood cells.

Mice are immunized with any of a chimeric capsomere vaccine and sera is collected as described above in Example 4. As positive controls, HPV16 L1 virus like particles (VLPs) and bovine PV1 (BPV) L1 VLPs are assayed in parallel with a chimeric capsomere preparation. To establish a positive baseline, the HPV16 or BPV1 VLPs are first incubated with or without sera collected from immunized mice after which red blood cells are added. The extent to which preincubation with mouse sera inhibits red blood cell hemagglutination is an indication of the neutralizing capacity of the mouse sera. The experiments are then repeated using chimeric capsomeres in order to determine the neutralizing effect of the mouse sera on the vaccine. A brief protocol for the hemagglutination inhibition assay is described below.

One hundred microliters of heparin (1000 usp units/ml) are added to 1 ml fresh mouse blood. Red blood cells are washed three times with PBS followed by centrifugation and resuspension in a volume of 10 ml. Next, erythrocytes are resuspended in 0.5 ml PBS and stored at 4°C for up to three days. For the hemagglutinin assay, 70 μ l of the suspension is used per well on a 96-well plate.

Chimeric capsomere aliquots from CsCl gradients are dialyzed for one hour against 10 mM Hepes (pH 7.5) and 100 μ l of two-fold serial dilutions in PBS are added to mouse erythrocytes in round-bottom 96-well microtiter plates which are further incubated for 3-16 hours at 4°C. For hemagglutination inhibition, capsomeres are incubated with dilutions of antibodies in PBS for 60 minutes at room temperature and then added to the erythrocytes. The level of erythrocyte hemagglutination, and therefore the presence of neutralizing antibodies, is determined by standard methods.

In preliminary results, mouse sera generated against chimeric capsomeres comprising HPV16L1 Δ C protein in association with E7 amino acid residues 1-98 was observed to inhibit hemagglutination by HPV16 VLPs, but not

by BPV VLPs. The mouse sera was therefore positive for neutralizing antibodies against the human VLPs and this differential neutralization was most likely the result of antibody specificity for epitopes against which the antibodies were raised.

Numerous modifications and variations in the invention as set forth in the above illustrative examples are expected to occur to those skilled in the art. Consequently only such limitations as appear in the appended claims should be placed on the invention.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(ii) TITLE OF INVENTION: Papilloma Virus Capsomere Vaccine Formulations and Methods of Use

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
(B) STREET: 233 South Wacker Drive, 6300 Sears Tower
(C) CITY: Chicago
(D) STATE: Illinois
(E) COUNTRY: United States of America
(F) ZIP: 60606-6402

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Williams Jr., Joseph A.
(B) REGISTRATION NUMBER: 38,659
(C) REFERENCE/DOCKET NUMBER: 27013/34028

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 312-474-6300
(B) TELEFAX: 312-474-0448

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1518 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..1518

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG	TCT	CTT	TGG	CTG	CCT	AGT	GAG	GCC	ACT	GTC	TAC	TTG	CCT	CCT	GTC	48
Met	Ser	Leu	Trp	Leu	Pro	Ser	Glu	Ala	Thr	Val	Tyr	Leu	Pro	Pro	Val	
1				5						10					15	
CCA	GTA	TCT	AAG	GTT	GTA	AGC	ACG	GAT	GAA	TAT	GTT	GCA	CGC	ACA	AAC	96
Pro	Val	Ser	Lys	Val	Val	Ser	Thr	Asp	Glu	Tyr	Val	Ala	Arg	Thr	Asn	
				20				25					30			

ATA	TAT	TAT	CAT	GCA	GGA	ACA	TCC	AGA	CTA	CTT	GCA	GTT	GGA	CAT	CCC	144
Ile	Tyr	Tyr	His	Ala	Gly	Thr	Ser	Arg	Leu	Leu	Ala	Val	Gly	His	Pro	
		35					40					45				
TAT	TTT	CCT	ATT	AAA	AAA	CCT	AAC	AAT	AAC	AAA	ATA	TTA	GTT	CCT	AAA	192
Tyr	Phe	Pro	Ile	Lys	Lys	Pro	Asn	Asn	Asn	Lys	Ile	Leu	Val	Pro	Lys	
		50				55					60					
GTA	TCA	GGA	TTA	CAA	TAC	AGG	GTA	TTT	AGA	ATA	CAT	TTA	CCT	GAC	CCC	240
Val	Ser	Gly	Leu	Gln	Tyr	Arg	Val	Phe	Arg	Ile	His	Leu	Pro	Asp	Pro	
		65			70					75					80	
AAT	AAG	TTT	GGT	TTT	CCT	GAC	ACC	TCA	TTT	TAT	AAT	CCA	GAT	ACA	CAG	288
Asn	Lys	Phe	Gly	Phe	Pro	Asp	Thr	Ser	Phe	Tyr	Asn	Pro	Asp	Thr	Gln	
				85					90					95		
CGG	CTG	GTT	TGG	GCC	TGT	GTA	GGT	GTT	GAG	GTA	GGT	CGT	GGT	CAG	CCA	336
Arg	Leu	Val	Trp	Ala	Cys	Val	Gly	Val	Glu	Val	Gly	Arg	Gly	Gln	Pro	
			100					105					110			
TTA	GGT	GTG	GGC	ATT	AGT	GGC	CAT	CCT	TTA	TTA	AAT	AAA	TTG	GAT	GAC	384
Leu	Gly	Val	Gly	Ile	Ser	Gly	His	Pro	Leu	Leu	Asn	Lys	Leu	Asp	Asp	
		115				120						125				
ACA	GAA	AAT	GCT	AGT	GCT	TAT	GCA	AAT	GCA	GGT	GTG	GAT	AAT	AGA		432
Thr	Glu	Asn	Ala	Ser	Ala	Tyr	Ala	Ala	Asn	Ala	Gly	Val	Asp	Asn	Arg	
		130				135					140					
GAA	TGT	ATA	TCT	ATG	GAT	TAC	AAA	CAA	ACA	CAA	TTG	TGT	TTA	ATT	GGT	480
Glu	Cys	Ile	Ser	Met	Asp	Tyr	Lys	Gln	Thr	Gln	Leu	Cys	Leu	Ile	Gly	
		145			150					155					160	
TGC	AAA	CCA	CCT	ATA	GGG	GAA	CAC	TGG	GGC	AAA	GGA	TCC	CCA	TGT	ACC	528
Cys	Lys	Pro	Pro	Ile	Gly	Glu	His	Trp	Gly	Lys	Gly	Ser	Pro	Cys	Thr	
				165				170						175		
AAT	GTT	GCA	GTA	AAT	CCA	GGT	GAT	TGT	CCA	CCA	TTA	GAG	TTA	ATA	AAC	576
Asn	Val	Ala	Val	Asn	Pro	Gly	Asp	Cys	Pro	Pro	Leu	Glu	Leu	Ile	Asn	
			180					185					190			
ACA	GTT	ATT	CAG	GAT	GGT	GAT	ATG	GTT	GAT	ACT	GGC	TTT	GGT	GCT	ATG	624
Thr	Val	Ile	Gln	Asp	Gly	Asp	Met	Val	Asp	Thr	Gly	Phe	Gly	Ala	Met	
		195					200					205				
GAC	TTT	ACT	ACA	TTA	CAG	GCT	AAC	AAA	AGT	GAA	GTT	CCA	CTG	GAT	ATT	672
Asp	Phe	Thr	Thr	Leu	Gln	Ala	Asn	Lys	Ser	Glu	Val	Pro	Leu	Asp	Ile	
		210				215					220					
TGT	ACA	TCT	ATT	TGC	AAA	TAT	CCA	GAT	TAT	ATT	AAA	ATG	GTG	TCA	GAA	720
Cys	Thr	Ser	Ile	Cys	Lys	Tyr	Pro	Asp	Tyr	Ile	Lys	Met	Val	Ser	Glu	
		225			230					235				240		
CCA	TAT	GGC	GAC	AGC	TTA	TTT	TTT	TAT	TTA	CGA	AGG	GAA	CAA	ATG	TTT	768
Pro	Tyr	Gly	Asp	Ser	Leu	Phe	Phe	Tyr	Ala	Arg	Arg	Glu	Gln	Met	Phe	
				245					250					255		
GTT	AGA	CAT	TTA	TTT	AAT	AGG	GCT	GGT	GCT	GTT	GGT	GAA	AAT	GTA	CCA	816
Val	Arg	His	Leu	Phe	Asn	Arg	Ala	Gly	Ala	Val	Gly	Glu	Asn	Val	Pro	
			260				265					270				
GAC	GAT	TTA	TAC	ATT	AAA	GGC	TCT	GGG	TCT	ACT	GCA	AAT	TTA	GCC	AGT	864
Asp	Asp	Leu	Tyr	Ile	Lys	Gly	Ser	Gly	Ser	Thr	Ala	Asn	Leu	Ala	Ser	
		275				280						285				

TCA AAT TAT TTT CCT ACA CCT AGT GGT TCT ATG GTT ACC TCT GAT GCC Ser Asn Tyr Phe Pro Thr Pro Ser Gly Ser Met Val Thr Ser Asp Ala 290 295 300	912
CAA ATA TTC AAT AAA CCT TAT TGG TTA CAA CGA GCA CAG GGC CAC AAT Gln Ile Phe Asn Lys Pro Tyr Trp Leu Gln Arg Ala Gln Gly His Asn 305 310 315 320	960
AAT GGC ATT TGT TGG GGT AAC CAA CTA TTT GTT ACT GTT GTT GAT ACT Asn Gly Ile Cys Trp Gly Asn Gln Leu Phe Val Thr Val Val Asp Thr 325 330 335	1008
ACA CGC AGT ACA AAT ATG TCA TTA TGT GCT GCC ATA TCT ACT TCA GAA Thr Arg Ser Thr Asn Met Ser Leu Cys Ala Ala Ile Ser Thr Ser Ser Glu 340 345 350	1056
ACT ACA TAT AAA AAT ACT AAC TTT AAG GAG TAC CTA CGA CAT GGG GAG Thr Thr Tyr Lys Asn Thr Asn Phe Lys Glu Tyr Leu Arg His Gly Glu 355 360 365	1104
GAA TAT GAT TTA CAG TTT ATT TTT CAA CTG TGC AAA ATA ACC TTA ACT Glu Tyr Asp Leu Gln Phe Ile Phe Gln Leu Cys Lys Ile Thr Leu Thr 370 375 380	1152
GCA GAC GTT ATG ACA TAC ATA CAT TCT ATG AAT TCC ACT ATT TTG GAG Ala Asp Val Met Thr Thr Ile His Ser Met Asn Ser Thr Ile Leu Glu 385 390 395 400	1200
GAC TGG AAT TTT GGT CTA CAA CCT CCC CCA GGA GGC ACA CTA GAA GAT Asp Trp Asn Phe Gly Leu Gln Pro Pro Pro Gly Gly Thr Leu Glu Asp 405 410 415	1248
ACT TAT AGG TTT GTA ACC TCC CAG GCA ATT GCT TGT CAA AAA CAT ACA Thr Tyr Arg Phe Val Thr Ser Gln Ala Ile Ala Cys Gln Lys His Thr 420 425 430	1296
CCT CCA GCA CCT AAA GAA GAT CCC CTT AAA AAA TAC ACT TTT TGG GAA Pro Pro Ala Pro Lys Glu Asp Pro Leu Lys Lys Tyr Thr Phe Trp Glu 435 440 445	1344
GTA AAT TTA AAG GAA AAG TTT TCT GCA GAC CTA GAT CAG TTT CCT TTA Val Asn Leu Lys Glu Lys Phe Ser Ala Asp Leu Asp Gln Phe Pro Leu 450 455 460	1392
GGA CGC AAA TTT TTA CTA CAA GCA GGA TTG AAG GCC AAA CCA AAA TTT Gly Arg Lys Phe Leu Leu Gln Ala Gly Leu Lys Ala Lys Pro Lys Phe 465 470 475 480	1440
ACA TTA GGA AAA CGA AAA GCT ACA CCC ACC ACC TCA TCT ACC TCT ACA Thr Leu Gly Lys Arg Lys Ala Thr Pro Thr Thr Ser Ser Thr Ser Thr 485 490 495	1488
ACT GCT AAA CGC AAA AAA CGT AAG CTG TAA Thr Ala Lys Arg Lys Lys Arg Lys Leu * 500 505	1518

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 506 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Ser Leu Trp Leu Pro Ser Glu Ala Thr Val Tyr Leu Pro Pro Val
 1           5           10           15
Pro Val Ser Lys Val Val Ser Thr Asp Glu Tyr Val Ala Arg Thr Asn
          20           25           30
Ile Tyr Tyr His Ala Gly Thr Ser Arg Leu Leu Ala Val Gly His Pro
      35           40           45
Tyr Phe Pro Ile Lys Lys Pro Asn Asn Asn Lys Ile Leu Val Pro Lys
      50           55           60
Val Ser Gly Leu Gln Tyr Arg Val Phe Arg Ile His Leu Pro Asp Pro
      65           70           75           80
Asn Lys Phe Gly Phe Pro Asp Thr Ser Phe Tyr Asn Pro Asp Thr Gln
          85           90           95
Arg Leu Val Trp Ala Cys Val Gly Val Glu Val Gly Arg Gly Gln Pro
      100          105          110
Leu Gly Val Gly Ile Ser Gly His Pro Leu Leu Asn Lys Leu Asp Asp
      115          120          125
Thr Glu Asn Ala Ser Ala Tyr Ala Ala Asn Ala Gly Val Asp Asn Arg
      130          135          140
Glu Cys Ile Ser Met Asp Tyr Lys Gln Thr Gln Leu Cys Leu Ile Gly
      145          150          155          160
Cys Lys Pro Pro Ile Gly Glu His Trp Gly Lys Gly Ser Pro Cys Thr
      165          170          175
Asn Val Ala Val Asn Pro Gly Asp Cys Pro Pro Leu Glu Leu Ile Asn
      180          185          190
Thr Val Ile Gln Asp Gly Asp Met Val Asp Thr Gly Phe Gly Ala Met
      195          200          205
Asp Phe Thr Thr Leu Gln Ala Asn Lys Ser Glu Val Pro Leu Asp Ile
      210          215          220
Cys Thr Ser Ile Cys Lys Tyr Pro Asp Tyr Ile Lys Met Val Ser Glu
      225          230          235          240
Pro Tyr Gly Asp Ser Leu Phe Phe Tyr Leu Arg Arg Glu Gln Met Phe
      245          250          255
Val Arg His Leu Phe Asn Arg Ala Gly Ala Val Gly Glu Asn Val Pro
      260          265          270
Asp Asp Leu Tyr Ile Lys Gly Ser Gly Ser Thr Ala Asn Leu Ala Ser
      275          280          285
Ser Asn Tyr Phe Pro Thr Pro Ser Gly Ser Met Val Thr Ser Asp Ala
      290          295          300

```

Gln Ile Phe Asn Lys Pro Tyr Trp Leu Gln Arg Ala Gln Gly His Asn
305 310 315 320

Asn Gly Ile Cys Trp Gly Asn Gln Leu Phe Val Thr Val Val Asp Thr
325 330 335

Thr Arg Ser Thr Asn Met Ser Leu Cys Ala Ala Ile Ser Thr Ser Glu
340 345 350

Thr Thr Tyr Lys Asn Thr Asn Phe Lys Glu Tyr Leu Arg His Gly Glu
355 360 365

Glu Tyr Asp Leu Gln Phe Ile Phe Gln Leu Cys Lys Ile Thr Leu Thr
370 375 380

Ala Asp Val Met Thr Tyr Ile His Ser Met Asn Ser Thr Ile Leu Glu
385 390 395 400

Asp Trp Asn Phe Gly Leu Gln Pro Pro Pro Gly Gly Thr Leu Glu Asp
405 410 415

Thr Tyr Arg Phe Val Thr Ser Gln Ala Ile Ala Cys Gln Lys His Thr
420 425 430

Pro Pro Ala Pro Lys Glu Asp Pro Leu Lys Lys Tyr Thr Phe Trp Glu
435 440 445

Val Asn Leu Lys Glu Lys Phe Ser Ala Asp Leu Asp Gln Phe Pro Leu
450 455 460

Gly Arg Lys Phe Leu Leu Gln Ala Gly Leu Lys Ala Lys Pro Lys Phe
465 470 475 480

Thr Leu Gly Lys Arg Lys Ala Thr Pro Thr Thr Ser Ser Thr Ser
485 490 495

Thr Ala Lys Arg Lys Lys Arg Lys Leu *
500 505

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 297 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..297

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG CAT GGA GAT ACA CCT ACA TTG CAT GAA TAT ATG TTA GAT TTG CAA	48
Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln	
1 5 10 15	
CCA GAG ACA ACT GAT CTC TAC TGT TAT GAG CAA TTA AAT GAC AGC TCA	96
Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser	
20 25 30	

GAG	GAG	GAG	GAT	GAA	ATA	GAT	GGT	CCA	GCT	GGA	CAA	GCA	GAA	CCG	GAC	144
Glu	Glu	Glu	Asp	Glu	Ile	Asp	Gly	Pro	Ala	Gly	Gln	Ala	Glu	Pro	Asp	
		35					40					45				
AGA	GCC	CAT	TAC	AAT	ATT	GTA	ACC	TTT	TGT	TGC	AAG	TGT	GAC	TCT	ACG	192
Arg	Ala	His	Tyr	Asn	Ile	Val	Thr	Phe	Cys	Cys	Lys	Cys	Asp	Ser	Thr	
		50				55					60					
CTT	CGG	TTG	TGC	GTA	CAA	AGC	ACA	CAC	GTA	GAC	ATT	CGT	ACT	TTG	GAA	240
Leu	Arg	Leu	Cys	Val	Gln	Ser	Thr	His	Val	Asp	Ile	Arg	Thr	Leu	Glu	
		65			70					75					80	
GAC	CTG	TTA	ATG	GGC	ACA	CTA	GGA	ATT	GTG	TGC	CCC	ATC	TGT	TCT	CAG	288
Asp	Leu	Leu	Met	Gly	Thr	Leu	Gly	Ile	Val	Cys	Pro	Ile	Cys	Ser	Gln	
				85					90					95		
AAA	CCA	TAA														297
Lys	Pro	*														

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 98 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	His	Gly	Asp	Thr	Pro	Thr	Leu	His	Glu	Tyr	Met	Leu	Asp	Leu	Gln	
1					5				10					15		
Pro	Glu	Thr	Thr	Asp	Leu	Tyr	Cys	Tyr	Glu	Gln	Leu	Asn	Asp	Ser	Ser	
			20					25					30			
Glu	Glu	Glu	Asp	Glu	Ile	Asp	Gly	Pro	Ala	Gly	Gln	Ala	Glu	Pro	Asp	
		35					40					45				
Arg	Ala	His	Tyr	Asn	Ile	Val	Thr	Phe	Cys	Cys	Lys	Cys	Asp	Ser	Thr	
		50			55						60					
Leu	Arg	Leu	Cys	Val	Gln	Ser	Thr	His	Val	Asp	Ile	Arg	Thr	Leu	Glu	
		65			70					75				80		
Asp	Leu	Leu	Met	Gly	Thr	Leu	Gly	Ile	Val	Cys	Pro	Ile	Cys	Ser	Gln	
			85					90						95		
Lys	Pro	*														

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCCCGATATC GCCTTTAATG TATAAATCGT CTGG

34

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCCCGATATC TCAAATTATT TTCCTACACC TAGTG

35

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAAGATATCT TGTAATAAAA ATTGCGTCC TAAAGGAAAC

40

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 44 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAAGATATCT AATCTACCTC TACAACTGCT AAACGCAAAA AAGC

44

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AAAAGATATC ATGCATGGAG ATACACCTAC ATTGC

35

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 34 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTIGATATC GGCTCTGTCC GGTTCTGCTT GTCC

34

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 44 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTTIGATATC CTGCAACAA AAGGTTACAA TATTGTAATG GGCC

44

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AAAAGATATC TGGTTTCTGA GAACAGATGG GGCAC

35

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTTIGATATC GATTATGAGC AATTAAATGA CAGCTCAG

38

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTTTGATATC GTCTACGTGT GTGCTTGTGA CGCAC

35

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TTTATCGATA TCGGTCCAGC TGGACAAGCA GAACCGGAC

39

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 39 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TTTTGATATC GATGCCCAT TACAATTGT AACCTTTTG

39

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 294 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ATG AGT CTT CTA ACC GAG GTC GAA ACG CTT ACC AGA AAC GGA TGG GAG 48
Met Ser Leu Leu Thr Glu Val Glu Thr Arg Asn Gly Trp Glu
1 5 10 15

TGC AAA TGC AGC GAT TCA AGT GAT CCT CTC ATT ATC GCA GCG AGT ATC 96
Cys Lys Cys Ser Asp Ser Ser Asp Pro Leu Ile Ile Ala Ala Ser Ile
20 25 30

ATT GGG ATC TTG CAC TTG ATA TTG TGG ATT TTT TAT CGT CTT TTC TTC 144
Ile Gly Ile Leu His Leu Ile Leu Trp Ile Phe Tyr Arg Leu Phe Phe
35 40 45

AAA TGC ATT TAT CGT CGC CTT AAA TAC GGT TTG AAA AGA GGG CCT TCT 192
Lys Cys Ile Tyr Arg Arg Leu Lys Tyr Gly Leu Lys Arg Gly Pro Ser
50 55 60

ACG	GAA	GGA	GCG	CCT	GAG	TCT	ATG	AGG	GAA	GAA	TAT	CGG	CAG	GAA	CAG	240
Thr	Glu	Gly	Ala	Pro	Glu	Ser	Met	Arg	Glu	Glu	Tyr	Arg	Gln	Glu	Gln	
65					70					75					80	
CAG	AGT	GCT	GTG	GAT	GTT	GAC	GAT	GTT	CAT	TTT	GTC	AAC	ATA	GAG	CTG	288
Gln	Ser	Ala	Val	Asp	Val	Asp	Asp	Val	His	Phe	Val	Asn	Ile	Glu	Leu	
			85						90					95		
GAG	TAA															294
Glu	*															

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 97 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met	Ser	Leu	Leu	Thr	Glu	Val	Glu	Thr	Leu	Thr	Arg	Asn	Gly	Trp	Glu	
1				5				10					15			
Cys	Lys	Cys	Ser	Asp	Ser	Ser	Asp	Pro	Leu	Ile	Ile	Ala	Ala	Ser	Ile	
			20					25					30			
Ile	Gly	Ile	Leu	His	Leu	Ile	Leu	Trp	Ile	Phe	Tyr	Arg	Leu	Phe	Phe	
			35				40					45				
Lys	Cys	Ile	Tyr	Arg	Arg	Leu	Lys	Tyr	Gly	Leu	Lys	Arg	Gly	Pro	Ser	
			50			55					60					
Thr	Glu	Gly	Ala	Pro	Glu	Ser	Met	Arg	Glu	Glu	Tyr	Arg	Gln	Glu	Gln	
65				70					75					80		
Gln	Ser	Ala	Val	Asp	Val	Asp	Asp	Val	His	Phe	Val	Asn	Ile	Glu	Leu	
			85					90						95		

Glu *

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 40 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TTTGTGATATC GATATGGAAT GGCTAAAGAC AAGACCAATC

40

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTTGTGATATC GTTGTTTGGA TCCCCATTCC CATTG 35

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTTATGACAT ACATACATTC TATG 24

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCATGCATTC CTGCTTGTAG TAAAAATTG CGTCC 35

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTACAAGCAG GAATGCATGG AGATACACC 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 36 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CATCTGAAGC TTAGTAATGG GCTCTGTCCG GTTCTG 36

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CATCTGAAGC TTATCAATAT TGTAATGGGC TCTGTCCG

38

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 54 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CATCTGAAGC TTACTTGCAA CAAAAGGTTA CAATATTGTA ATGGGCTCTG TCCG

54

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 69 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CATCTGAAGC TTAAAGCGTA GAGTCACACT TGCAACAAAA GGTTACAATA TTGTAATGGG

60

CTCTGTCCG

69

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 47 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CATCTGAAGC TTATTGTACG CACAACCGAA GCGTAGAGTC ACAC TTG

47